Genetic diversity of *Crotalaria* germplasm assessed through phylogenetic analysis of EST-SSR markers

M.L. Wang, J.A. Mosjidis, J.B. Morris, R.E. Dean, T.M. Jenkins, and G.A. Pederson

Abstract: The genetic diversity of the genus *Crotalaria* is unknown even though many species in this genus are economically valuable. We report the first study in which polymorphic expressed sequence tag-simple sequence repeat (EST-SSR) markers derived from *Medicago* and soybean were used to assess the genetic diversity of the *Crotalaria* germplasm collection. This collection consisted of 26 accessions representing 4 morphologically characterized species. Phylogenetic analysis partitioned accessions into 4 main groups generally along species lines and revealed that 2 accessions were incorrectly identified as *Crotalaria juncea* and *Crotalaria spectabilis* instead of *Crotalaria retusa*. Morphological re-examination confirmed that these 2 accessions were misclassified during curation or conservation and were indeed *C. retusa*. Some amplicons from *Crotalaria* were sequenced and their sequences showed a high similarity (89% sequence identity) to *Medicago truncatula* from which the EST-SSR primers were designed; however, the SSRs were completely deleted in *Crotalaria*. Highly distinguishing markers or more sequences are required to further classify accessions within *C. juncea*.

Key words: Crotalaria germplasm, EST-SSR, genetic diversity, phylogeny.

Résumé : La diversité génétique au sein du genre *Crotalaria* est inconnue malgré le fait que plusieurs espèces aient une utilité économique. Les auteurs rapportent la première étude où des microsatellites de régions codantes (EST-SSR) polymorphes dérivés de *Medicago* ou du soya ont été employés pour examiner une collection de germoplasme de *Crotalaria*. Cette collection comprenait 26 accessions représentant quatre espèces caractérisées sur le plan morphologique. Une analyse phylogénétique a séparé les espèces en quatre groupes principaux essentiellement en conformité avec les espèces. Deux accessions étaient erronément classées au sein de *Crotalaria juncea* et *Crotalaria spectabilis* au lieu de *Crotalaria retusa*. Un nouvel examen morphologique a confirmé que ces espèces avaient été mal classifiées au moment de leur saisie ou de leur conservation et qu'elles appartenaient véritablement au *C. retusa*. Certains amplicons de *Crotalaria* ont été séquencés et leurs séquences montraient une grande similarité (89 % d'identité) avec les séquences de *Medicago truncatula* d'où provenaient les amorces microsatellites. Cependant, la portion microsatellite était absente chez *Crotalaria*. Des marqueurs offrant une plus grande discrimination et plus de séquences seront nécessaires afin de classer plus finement les accessions au sein de *C. juncea*.

Mots clés : germoplasme du genre *Crotalaria*, EST-SSR, diversité génétique, phylogénie.

[Traduit par la Rédaction]

Introduction

The genus *Crotalaria*, which belongs to the family Papilionaceae, the subfamily Papilionoideae, and the tribe Crotalarieae (van Wyk and Schutte 1995), contains approximately 550 species dispersed throughout the temperate, subtropical, and tropical regions of the world. They are used as a source of fibers, silage, and green manure for nitrogen fixation (Dempsey 1975; Ramos et al. 2001); as cover crops for control of weeds, nematodes, and soil erosion; and as ornamentals (Miller 1967). Some *Crotalaria* species can be used as forage for horses and cattle owing to the large amounts of water-soluble gums and proteins in their seeds (Purseglove 1981; Pandey and Srivastava 1990). Despite this commercial value, the genetic diversity of *Crotalaria* species germplasm has not been assessed.

Received 14 October 2005. Accepted 14 February 2006. Published on the NRC Research Press Web site at http://genome.nrc.ca on 30 June 2006.

Corresponding Editor: J.P. Gustafson.

M.L. Wang,[1] J.B. Morris, and G.A. Pederson. USDA–ARS, PGRCU, 1109 Experiment Street, Griffin, GA 30223, USA.
J.A. Mosjidis. Department of Agronomy and Soils, Auburn University, 202 Funchess Hall, Auburn, AL 36849, USA.
R.E. Dean. University of Georgia, Plant Genetic Resources Conservation Unit, 1109 Experiment Street, Griffin, GA 30223, USA.
T.M. Jenkins. Department of Entomology, University of Georgia, 1109 Experiment Street, Griffin, GA 30223, USA.

[1]Corresponding author (e-mail: mwang@ars-grin.gov).

 doi:10.1139/G06-027

Table 1. Sunn hemp accessions

PI No.	Species	Identifer	Collection location
189043-1	*Crotalaria juncea* L.	Auburn, Ala., tissue	Nigeria
189043-2	*Crotalaria juncea* L.	Griffin, Ga., seeds	Nigeria
207657	*Crotalaria juncea* L.	n.a.	Sir Lanka
219717	*Crotalaria juncea* L.	n.a.	Myanmar
234771	*Crotalaria juncea* L.	n.a.	Nigeria
250485	*Crotalaria juncea* L.	K679	India
250486	*Crotalaria juncea* L.	K680	India
250487	*Crotalaria juncea* L.	K681	India
274948	*Crotalaria juncea* L.	n.a.	Guadeloupe
295851	*Crotalaria juncea* L.	n.a.	Brazil
314239	*Crotalaria juncea* L.	COL NO 524	Former Soviet Union
322377	*Crotalaria juncea* L.	IRI 2473	Brazil
337080	*Crotalaria juncea* L.	n.a.	Brazil
346297	*Crotalaria juncea* L.	n.a.	India
391567	*Crotalaria juncea* L.	T'ai-yang-ma	South Africa
426626	*Crotalaria juncea* L.	Sanni K-98	Pakistan
468956	*Crotalaria juncea* L.	'Tropic Sun'	USA
561720	*Crotalaria juncea* L.	IAC-1	Brazil
198000	*Crotalaria pallida* Aiton	No. 1	Brazil
186304	*Crotalaria pallida* Aiton	n.a.	Australia
274951	*Crotalaria retusa* L.	n.a.	Senegal
247128	*Crotalaria retusa* L.	n.a.	Brazil
186445	*Crotalaria retusa* L.	n.a.	Cameroon
Seral	*Crotalaria spectablis* Roth	'Georgia'	USA
322403	*Crotalaria spectablis* Roth	IRI 1788	Brazil
217908	*Crotalaria spectablis* Roth	n.a.	India
240413	*Crotalaria spectablis* Roth	n.a.	Australia

Note: PI, plant introduction; n.a., no data available.

Crotalaria juncea (haploid chromosome number $x = 8$), commonly called sunn hemp, and which originated in India (Purseglove 1981), is a fast-growing, cross-pollinated, diploid species rich in stem and seed dietary fiber (Morris and Kays 2005). Sunn hemp seeds contain various pyrrolizidine alkaloids such as junceine, riddelliine, senecionine, seneciphylline, and trichodesmine (Smith and Culvenor 1981), which can be toxic to animals and birds when ingested in sufficient amounts (Purseglove 1981). Additionally, Ji et al. (2005) analyzed sunn hemp seeds from 9 populations that originated in different parts of the world for several pyrrolizidine alkaloids and determined that these populations had small amounts of junceine and trichodesmine (measured pyrrolizidine alkaloids were 0.8 to 3.8 μmol/g). These values represented 1.4%–6.3% of the amount of alkaloids present in the seed of *Crotalaria spectabilis* Roth, a species known to be toxic to animals and birds. Furthermore, a recent study with chickens indicated that mortality rate was not affected by inclusion of sunn hemp seeds in the diet (Hess and Mosjidis, unpublished results), which supports the report that sunn hemp seeds were non-toxic to chicks when fed at 10 mg/g body mass (Williams and Molyneux 1987).

Transferred SSR markers have been used for assessment of genetic diversity and examination of phylogenetic relationships of plant germplasm (Wang et al. 2004; Barkley et al. 2005; Wang et al. 2005; Wang et al. 2006). We are not aware of any report, however, in which transferred expressed sequence tag simple sequence repeat (EST-SSR) markers were used to assess genetic diversity of *Crotalaria* germplasm, examine the phylogenetic relationships between *Crotalaria* species, and compare the sequence content of amplicons within and among *Crotalaria* species. The objectives of the study are, therefore, to assess the genetic diversity of the *Crotalaria* germplasm through phylogenetic analysis of EST-SSR markers, verify the species designation of each accession, and compare the sequence content of cross-species amplicons generated from the transferred EST-SSR markers.

Materials and methods

The *Crotalaria* germplasm collection is maintained by the United States Department of Agriculture, Agricultural Research Service (USDA–ARS), National Plant Germplasm System (NPGS), Plant Genetic Resources Conservation Unit (PGRCU) in Griffin, Ga. Twenty-six accessions (Table 1) were used in this experiment. Among them were 17 accessions from *Crotalaria juncea* L. (sunn hemp, cross-pollinated species); 2 accessions from *Crotalaria pallida* Aiton (self-pollinated species); 3 accessions from *Crotalaria retusa* L. (self-pollinated species); and 4 accessions from *Crotalaria spectabilis* Roth (self-pollinated species). The samples were collected from leaf tissue of plants either grown in Auburn, Ala., or in Griffin, Ga. PI 189043 was collected from both locations (-1 from Auburn, Ala., and -2 from Griffin, Ga.) and used as an internal control for classification. PI 468956 is a cultivar, *Crotalaria juncea* L.

'Tropic Sun', which is being used in the United States. A feral *Crotalaria* (collected on a roadside in Tifton, Ga., and tentatively classified as *C. spectabilis*) was also included. For morphological comparison, leaf-tip shape, plant height, and seed shape were observed and recorded in the Griffin greenhouse.

DNA was extracted from leaf tissue using an EZNA plant DNA kit from Omega Bio-Tek (Doraville, Ga.) and diluted to 10 ng/μL for PCR. Fifty-eight EST-SSR primers (54 from *Medicago truncatula* and 4 from *Glycine max* (L.), Table 2) were used for screening polymorphisms. Methods used for the PCR cycling programs and product separation on agarose gels followed those described by Wang et al. (2005).

Strong clear bands on the gel images were scored using bins and as either present (1) or absent (0). The data was entered into a binary matrix for analysis. A similarity matrix was created using the algorithm for generating a proportion of alleles shared matrix in MICROSAT (Minch et al. 1996; http://hpgl.stanford.edu/projects/microsat/microsat.html). One hundred bootstrap matrices were generated and a consensus tree was obtained using the neighbor-joining (NJ) and consense programs in the Phylip suite (version 3.6; Felsenstein 1993). The trees were then viewed and printed using Treeview (Page 1996). Macros were written in Excel (Microsoft, Redmond, Wash.), which generated a shared-distance matrix based on the proportion of alleles shared to verify tree consistency. The overall topology remained unchanged.

For sequencing cross-species amplicons, PCR products were checked on agarose gels. PCR products containing only a single band in the solution were digested with Exonuclease I (at a final concentration of 0.1 U/μL; catalogue No. 70073X, USB, Cleveland, Ohio) and dephosphorylated with Shrimp Alkaline Phosphatase (SAP; catalogue No. 70092, USB) at 37 ºC for 15 min and then at 80 ºC for 15 min to inactivate the enzymes. The treated PCR products were transferred into a spin column (Wizard minicolumns, PRA7211, Promega, Madison, Wis.) and centrifuged at 8900*g* for 1 min in an Eppendorf centrifuge 5415D. The collected solution was used as DNA template for the sequencing reaction. The amplicons were sequenced on an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Foster City, Calif.) at the University of Georgia sequencing facility. To avoid sequencing errors, an amplicon was sequenced twice in both directions. The sequences were edited using the computer program DNA Star (http://www.dnastar.com). The DNA sequences were then aligned using the CLUSTAL method (in DNA Star).

Results and discussion

Phylogenetic analysis

Of the 58 sets of primers, 28 (48%, Table 1) generated polymorphic amplicons (DNA fragments or bands on agarose gels). Some primers generated multiple amplicons and some primers generated single amplicons (Fig. 1). A total of 130 bands were detected with a mean number of 4.6 alleles/marker. These polymorphic primers (Table 1) were used to assess genetic diversity of 26 *Crotalaria* accessions. A dendrogram was generated based on these polymorphic banding patterns. The accessions were classified into 4 main groups (Fig. 2), supported by strong nodes (86%–100%) (Fig. 2). Group I (or *C. pallida* group) contained 2 accessions (PI 186304 and PI 198000) from *C. pallida*. Group II (or *C. juncea* group) contained 16 accessions (from PI 274948 to PI 250486), all of which were from *C. juncea*. Group III (or the *C. retusa* group) contained 5 accessions, 3 of which were from *C. retusa*, and 2 of which had been previously thought to belong to different species, 1 from *C. spectabilis* and 1 from *C. juncea*. Group IV (or the *C. spectabilis* group) contained 3 accessions, all of which were from *C. spectabilis*. Two derivatives (PI 189043-1 and PI 189043-2) from the same accession but that were requested from 2 different locations should be grouped together. They indeed grouped into a single internode (with a bootstrap value of 61, not shown in Fig. 2), but they were not identical. Overall, the dendrogram (Fig. 2) demonstrated that the genetic relatedness of *C. juncea* with *C. retusa* and *C. spectabilis* was much closer than that with *C. pallida*. *Crotalaria retusa* was also closely related to *C. spectabilis*.

Genetic variation among and within species

As seen from the dendrogram (Fig. 2) the groups (or clades) were species specific. Notice that a couple of species (labeled with a star) were apparently misclassified; however, upon further morphological re-examination, these were changed in accordance with the observations from the dendrogram. With the exception of *C. pallida* and accessions 337080 and 250486, each clade shared some 90% or more alleles. When compared outside their clades, this number dropped noticeably. Within species, 2 accessions (PI 186304 and PI 198000) from *C. pallida* were allocated to the same group (used as outgroup). However, these 2 accessions were very different genetically. Within the species *C. juncea*, most of the accessions could not be distinguished well because most of the bootstrap values were below 50%. Further distinguishing the accessions within *C. juncea* would require markers with a high distinguishing power or the sequencing of more amplicons. Interestingly, accession PI 189043 was clearly part of the *C. retusa* group, but this accession was also visibly separated from the other *C. retusa* accessions. Accession PI 322403 from *C. spectabilis* was also classified into the *C. retusa* group. However, this accession was very closely related to other *C. retusa* accessions (PI 186445, PI 247128 and PI 274951). For accessions within the species *C. spectabilis*, 3 accessions (feral, PI 217908 and PI 240413) were grouped together, but 2 accessions (PI 217908 and PI 240413) were more closely related to each other than to the feral plant.

Distinguishing power of transferred EST-SSR markers

Transferred EST-SSR markers from *Medicago* and soybean were better at detecting inter-specific rather than intra-specific polymorphism. This result was consistent with what was observed in the grass family (Wang et al. 2006). Marker AW584539, for example, generated multiple amplicons (Fig. 1) that facilitated species identification, but only 2 accessions (PI 198000 and PI 186304) from the same species *C. pallida* could be distinguished by these amplicons. The band patterns of accessions PI 189043 and PI 322403 were different from their designated species. Subsequent morphological comparison confirmed that these 2 accessions were misnamed and should be included in the *C. retusa* group.

Table 2. Primer information.

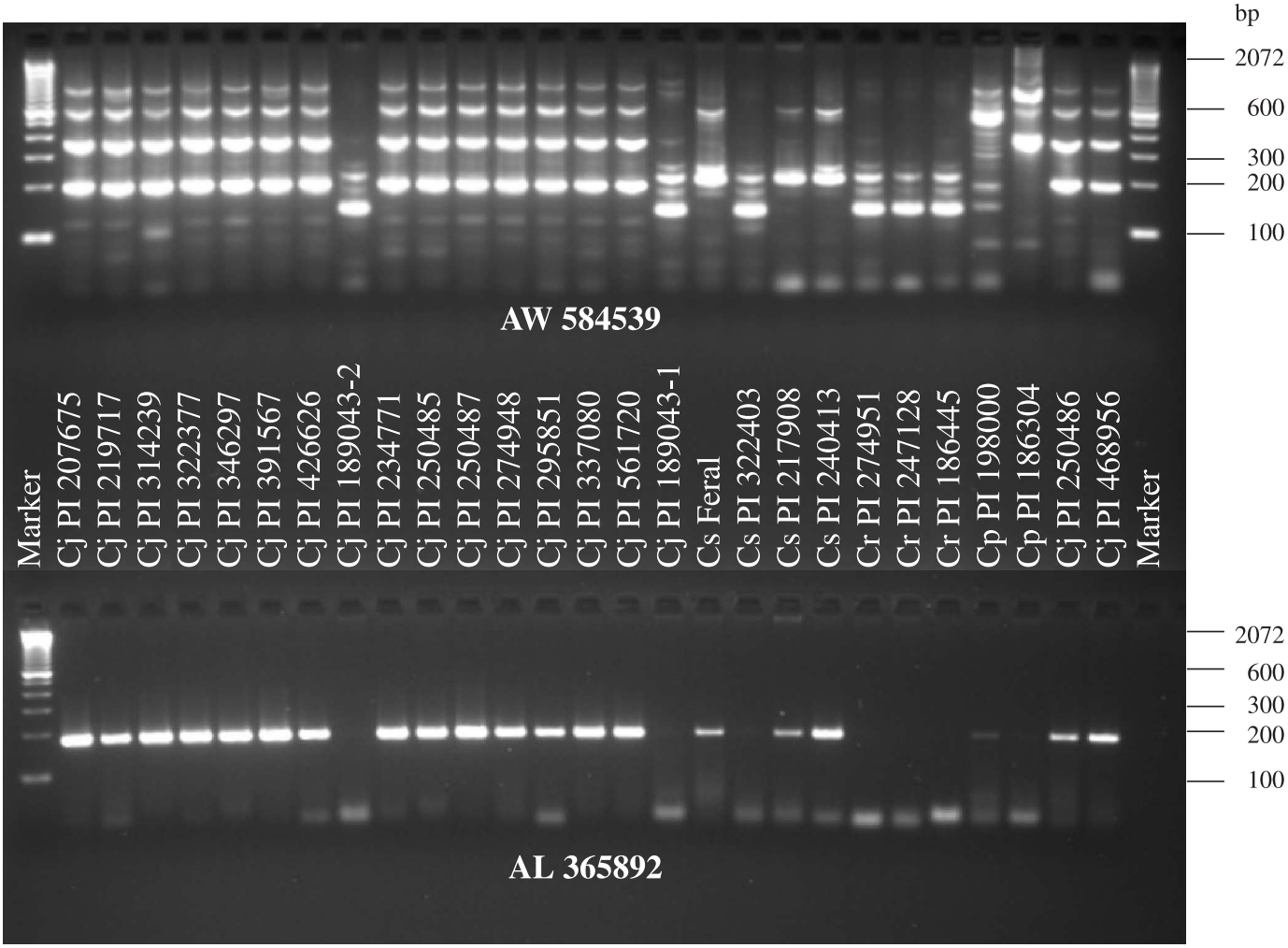
Name	Origin	Repeat	Forward primer (5′→3′)	Reverse primer (5′→3′)	Size (bp)
AL370549	*Medicago*	$(AC)_{11}$	CGTCCCGATATCGTCAACTT	CCACCACGACACATGTTACC	198
BF650979	*Medicago*	$(AT)_{28}$	TTGTGGAAGGAACAACTCTGG	GAAACCGGCATGATTAAGACA	179
AQ842128*	*Medicago*	$(TA)_{23}$	TCAATGCTGATGCCATTTTC	TCGCGTATTATAGCACAACACC	209
MSA13293	*Medicago*	$(CT)_{16}$	AGACAAGCGTGACACCCACT	TCCATCGCTCTCTCTTTTCC	209
AI974357*	*Medicago*	$(TC)_{25}$	TCTCAATTCCCCAACTTGCT	TCTCCTTCACCCATCTTTGC	183
AW256794	*Medicago*	$(TC)_{17}$	GTCATCGAAGGCCAAAACAC	GTTTGCGAGAAACACCGATT	192
AW560742	*Medicago*	$(TC)_{20}$	CTCCGTCCAACACATACCACT	TTCGATTCCGACTTCGATTT	198
AI737608	*Medicago*	$(GA)_{32}$	CGTCGTTTCAGAGGTCGAAG	AGCAACGTTGATTCCATGTG	199
AW584539*	*Medicago*	$(ACA)_{8}$	TTGATGGGCAATACATGTCG	GTTGAAGGAAGGTGGTGGTG	204
AW586959	*Medicago*	$(ACA)_{10}$	CGAGAATCATCGTAATTGGACA	CGAAGTTCAATGGCATCAGA	222
AW775229*	*Medicago*	$(AGC)_{8}$	TACTGGGGTGATGCAAGACA	CAATACCCAGAGGAGCAGCTA	222
AW684341	*Medicago*	$(ATC)_{9}$	ACCTGGATTGCTGTTTCCAC	CCCACTTGCAGCCTTCTCTA	203
AL365892*	*Medicago*	$(ATT)_{8}$	CCTCCACATAGCTGGTCGAT	GGTGTGTGCGGTTTAGGACT	216
BF649209*	*Medicago*	$(CCA)_{7}$	AAGAGGCGGAGAGTGAGGTT	GGTAAGAGAACGAGCGAGGA	200
AQ579641	*Medicago*	$(CTT)_{11}$	TTCACACGTTTGTGGACCTC	CCGACCTAACGGACCACTAC	230
AW685679	*Medicago*	$(GCC)_{5}$	ACCTCACCTCACCTCCCTTT	GATCATCTGGGTTTCGCAAG	197
AI974841*	*Medicago*	$(TCT)_{11}$	TCACCACCAAACCCCAAC	TGGCAATGCTACAAGCCTAA	169
AW688216*	*Medicago*	$(AGTG)_{9}$	CACGAGGGATTGTTGTTTGA	GGAGCAGTAGGGTTGCATCT	211
BI263393	*Medicago*	$(AGTG)_{9}$	TCGGCACGAGTTTGAGATAA	GGAGCAGTAGGGTTGCATCT	206
AW127626*	*Medicago*	$(GTTT)_{7}$	CATTTTGAAGGAAGGAAGAAGG	ATTTGGAAGCGGAATGTGAA	191
AW688861	*Medicago*	$(CAACT)_{7}$	TTGTTGTGTGGCTTCTTTGG	AAACCAACCACCTGTGTTGAC	195
AL369994*	*Medicago*	$(TGCGA)_{6}$	GGCGAAACCAATCAACATCT	CATCAGGTGCGTAGCATCAT	213
MtSSRNFAL18*	*Medicago*	$(TGTCA)_{5}$	TGTTTTTGCCAGTTTCTTGG	CACTCACGACCAACCACATT	200–210
MtSSRNFAL21	*Medicago*	$(AG)_{10}$	CGGTTGCACCAGAAAAACAT	TCTACGGCCCTTGAAGTCAC	184–204
MtSSRNFAL29	*Medicago*	$(TTA)_{8}$	TGGATGAATTTGGTGTAGGG	CAATACCCAGAGGAGCAGCTA	180–206
MtSSRNFAW10*	*Medicago*	$(TA)_{11}$	AGAAGGGAGCACAGATGACC	TCATTACCGAGGTTTCAAGGA	168–186
MtSSRNFAW114	*Medicago*	$(TTC)_{8}$	TGATAAGAACGAACAACGATAACC	GTCGGGAATAGCTCCGTCTT	190–234
MtSSRNFAW142*	*Medicago*	$(ATAG)_{5}$	AGGATCTAGCATGTTCTCAGATTT	AGGCCATGGAGGAGGACTA	198–224
MtSSRNFAW152*	*Medicago*	$(ATCCA)_{4}$	ACGGATTTCACTTCCATTCAAA	AGATCATGCAAGGCTTCAAAAC	194–316
MtSSRNFAW154*	*Medicago*	$(CAA)_{7}$	CCCTTCACCTTCTTCTTCACAC	TCAGTGCCAAAACAGATCTCC	700–800
MtSSRNFAW16*	*Medicago*	$(TTC)_{17}$	ATCGTCCCCACTGTGTCTTC	GTGGGGTTGGTGAGAGTGTT	142–175
MtSSRNFAW175	*Medicago*	$(ATC)_{7}$	CCCATAGGTAAAAACATCAAATGC	TTTTGGAGTCGGAGGAAGAA	194–232
MtSSRNFAW26	*Medicago*	$(GAAA)_{5}$	AGCTAGAACCTACCCTTGGAA	TAGGAGAACCGGCGATGTAG	159–186
MtSSRNFAW38*	*Medicago*	$(AAGA)_{5}$	ACCCAAATCTGTGCACGAAC	CAAATCCAAATGGCGTAACC	202–228
MtSSRNFAW50*	*Medicago*	$(AGC)_{8}$	TACTGGGGTGATGCAAGACA	TGGCAATGCTACAAGCCTAA	222–276
MtSSRNFAW55*	*Medicago*	$(CAAAA)_{4}$	GATTGGAACCGCTCAATCAT	GGGTCAAAGAAAAAGCCAAC	160–178

Table 2 (*concluded*).

Name	Origin	Repeat	Forward primer (5′→3′)	Reverse primer (5′→3′)	Size (bp)
MtSSRNFAW78*	*Medicago*	$(TTC)_9$	TCACTCACATAACTCCAAAGAACA	GGCTCCATGGTTCAGGTAAA	192–210
MtSSRNFAW81	*Medicago*	$(CAACT)_7$	TTGTTGTGTGGCTTCTTTGG	AAACCAACCACCTGTGTTGAC	164–190
MtSSRNFBE15*	*Medicago*	$(ATATA)_6$	TCACACTGCACAAGCATAACC	CGTGGTGGTCGGACTTATCT	202–230
MtSSRNFBF08	*Medicago*	$(CAT)_{10}$	CAGAAATTTCCATGCCAAAA	GGCGATTGTTCAAGCTGTTA	198–210
MtSSRNFBF103*	*Medicago*	$(CT)_{10}$	CCCTTTCACACACTCTTTCCTT	CGGTGTCGATAGATTGAACCTC	158–240
MtSSRNFBF20	*Medicago*	$(ATTC)_5$	GTTGGTGGTGGTCTTGGAAT	GCAGGAACGCTCTCAAACAT	230–312
MtSSRNFBF42	*Medicago*	$(TTC)_{17}$	ATCGTCCCCACTGTGTCTTC	TCGAATACGTCTTGGTCTTGG	153–183
MtSSRNFBF91*	*Medicago*	$(TC)_{10}$	CGTTTCGCGTTTCTCTTTCT	TGACGAATATCAGCCTCGGTA	170–238
MtSSRNFBG14	*Medicago*	$(ACC)_9$	GCTTTCCCTTCCACTCTACTCA	AAGGAGGTGCTGATTTGACG	222–238
MtSSRNFBG24	*Medicago*	$(TTAA)_5$	TCAATGATTGGGAAACCAAG	ATTGGGACAAAGGGGATAGG	145–202
MtSSRNFBG28*	*Medicago*	$(GGTTT)_5$	GAGCAAAGGGGTTTGTCTCA	GCAACTCCAGCTGCATCTTT	174–198
MtSSRNFBG37	*Medicago*	$(TTC)_{10}$	CAACCCCCAGCACATTTATT	GCAGGAACGCTCTCAAACAT	142–192
MtSSRNFBG46*	*Medicago*	$(ATATG)_4$	TGAGAAGAAAGGGATCATGGA	CGAACCAACTCACTCAAAAACA	148–248
MtSSRNFBG52	*Medicago*	$(TCTT)_6$	GGGAGAATTGCGAAGCATAA	TCGCATTAAAGGGAGATTTG	190–250
MtSSRNFBI14	*Medicago*	$(TTC)_{12}$	TCAACCGCATTTTTCTCTCC	TTACTGGTTCTTAATTTTATCCATCA	212–240
MtSSRNFBI15	*Medicago*	$(TAA)_7$	GGGAAAATTCAAACAAGGTGA	AGGTGCTGACGTTTGGAAAG	200–215
MtSSRNFBI49*	*Medicago*	$(TCT)_7$	AGCTCGCACACTTTCTCGTT	TGCAGCATCCACAAACTTTC	304–320
MtSSRNFBI52*	*Medicago*	$(CTC)_7$	CAACCCTTCCCTCTATCCTCTC	TCAGCTGTAGTTGGAGCAACAT	340–480
AG81	Soybean	$(AG)_{13}$	ATTTTCCAACTCGAATTGACC	TCATCAATCTCGACAAAGAATG	n.a.
BE801128	Soybean	$(CAA)_{13}$	GCGACAGTTCTCCACTCTTC	GCGCCCCTTATAGATTTGTAAC	172
AW508247	Soybean	$(CTT)_{10}$	GCGCCCAATCCCAATCTCAC	GCGAAGCCAATAAATGATAAAAATC	153
AW186493*	Soybean	$(CTT)_{13}$	GCGGTGATCCGTGAGATG	GCGGAAAGTAGCACCAAGAG	219

Note: n.a., not applicable.
*Markers were used for characterization of sunn hemp germplasm.

Fig. 1. Amplicons generated by PCR and separated by electrophoresis. Each well contains either 10 µL of molecular marker (100 bp ladder, 250 ng) or 12.5 µL of PCR products. The fragments were separated by electrophoresis on a 3% agarose gel. PCR products were generated with the primer pairs AW 584539 (upper panel) and AL 365892 (lower panel) from *Medicago*. Abbreviations used are as follows: Cj, *Crotalaria juncea*; Cp, *Crotalaria pallida*; Cr, *Crotalaria retusa*; Cs, *Crotalaria spectabilis*; bp, base pairs.

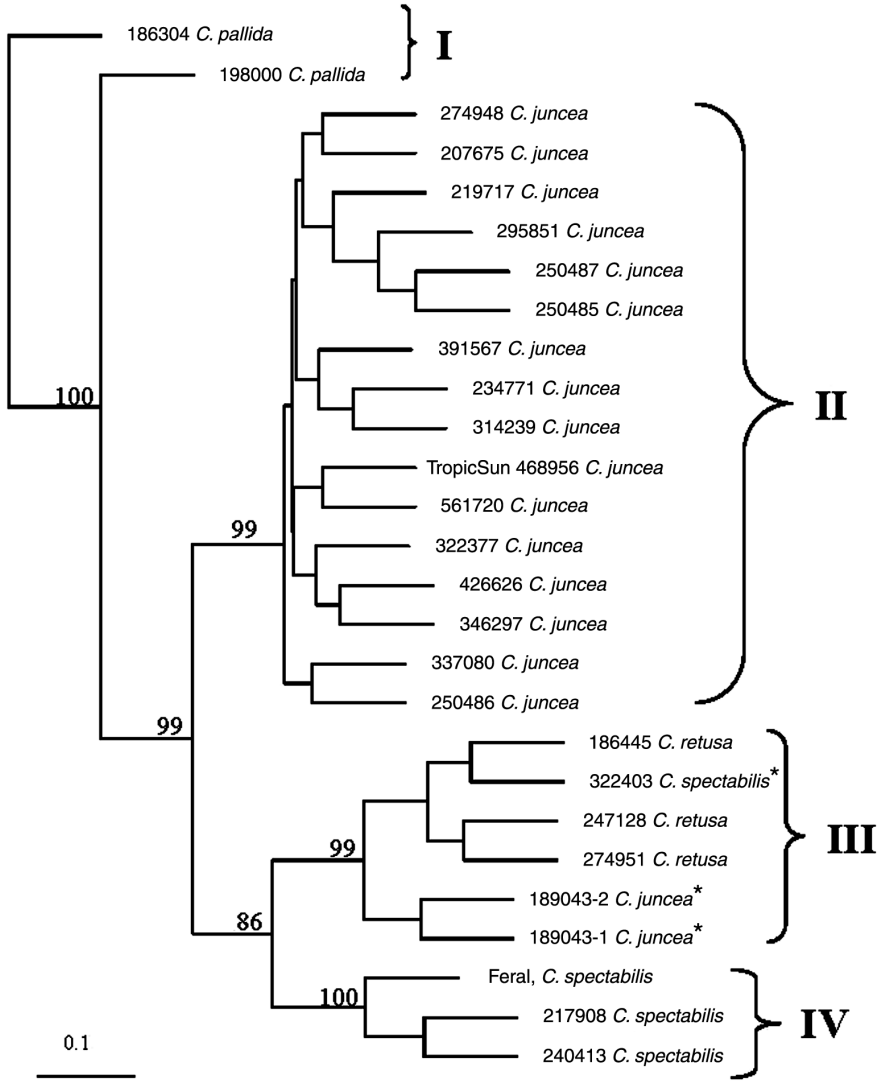


Some primers generated multiple amplicons that could be easily used to distinguish different species, but they were not necessarily powerful in distinguishing different accessions within the same species. To a certain extent, it seemed that multiple amplicons generated from the same set of primers could be used as fingerprints to distinguish accessions from different species (Fig. 1). However, the study showed that phylogenetic analysis of molecular data generated from EST-SSR markers was appropriate for determining the genetic diversity, taxonomy, and phylogeny (Avise 1994) of a *Crotalaria* collection.

Identification and reconfirmation of misnamed accessions

This study showed that phylogenetic analysis of molecular data can be effectively used to determine genetic variation as well as interspecific extant relationships. Accessions with a similar genetic composition can be classified into the same group. Similar phenotypes can usually be observed from accessions within the same group. The phylogenetic analysis suggests that accession PI 189043 should not be included in the species *C. juncea* and PI 322403 should not be included in the species *C. spectabilis*. Most likely these 2 accessions were misidentified during curation and conservation. Both of them were closely related to *C. retusa*. To confirm their genetic relationship with *C. retusa*, these 2 accessions and several others from *C. juncea*, *C. spectabilis*, and *C. retusa* were replanted in the greenhouse for morphological observation and comparison. At the young seedling stage the leaf-tip shapes of *C. juncea*, *C. retusa*, and *C. spectabilis* were pointed, spindle, and rounded, respectively (Fig. 3). Typically, *C.* juncea seeds are large (6–7 mm long and 4–4.75 mm wide), contorted on one surface, dull to lustrous or dull glossy, grayish olive green to olive gray, light olive brown to bluish purple; *C. spectabilis* seeds are 4 –5 mm long and 2.75–3.25 mm wide, dull glossy to high glossy, moderate olive brown to blackish brown; and *C. retusa* seeds are the smallest of the 3 species (3.25–3.5 mm long and 2.75–3 mm wide, dark orange yellow to dark red brown (Miller 1967). When comparing mature seeds, the seed size of *C. juncea* was bigger than *C. spectabilis* and had the seed colour of *C. juncea* was a lighter black. The seed color of

Fig. 2. Dendrogram of *Crotalaria* generated by neighbor-joining method. Bootstrapping was performed with 100 replicates and its value was labeled within each internode. The star represents the accessions misclassified by morphology during curation or conservation. Abbreviations used are the same as those in Fig. 1. Scale bar represents coefficient similarity.



C. retusa was golden yellow and very different from the former 2 species (Fig. 3). The plant morphology and seed shape of accessions PI 189043 and PI 322403 were more like *C. retusa*. Combining the genetic analysis and morphological comparison together, PI 189043 and PI 322403 were therefore reclassified as *C. retusa*.

Sequence comparison of amplicons generated from different species

Characterization and evaluation of plant germplasm from minor species by transferred EST-SSR markers from model species are mainly dependent on the phylogenetic relationship of the species, primer sequences, and the PCR conditions (PCR mixture and program) employed. At fixed PCR conditions, amplicons should be generated or not generated across species. Even though amplicons can be generated across species, it does not necessarily mean that the sequence content of amplicons from minor species and model species are similar or identical. Thus, since molecular characters were subjected to quantification under the hypothesis of monophy, the underlying molecular sequence had to be evaluated. To determine the sequence content, as an example, cross-species amplicons generated by EST-SSR marker AL365892 were sequenced. The size of cross-species amplicons was 183 bp. When these sequences were BLASTed against GenBank, the highest hit was a *Medicago* EST, AL365892, which contains SSRs and from which the primers used in this study were designed. In comparison with the original sequence from *Medicago*, all sequences were aligned, as shown in Fig. 4. The primer sequences matched exactly (underlined sequences) and the sequence identity between *Crotalaria* and *Medicago* was also very high (89%); however, the repeat sequences $(ATT)_8$ were completely missing in *Crotalaria* (bold sequences). Our result was consistent with some wheat and maize amplicons generated from rice SSRs that did not contain expected SSRs (Chen et al. 2002). In addition, there existed another 2 deletions (ATATTC and ATC, in italics) in the form of triple nucleotides. Interestingly, all of the deletions in this sequenced region were in the form of triple nucleotides. This

Fig. 3. Morphological comparison of some accessions from different species. Leaf-tip shape was photographed at the seedling stage and seed shape was photographed at maturation. Scale bar represents 10 cm.

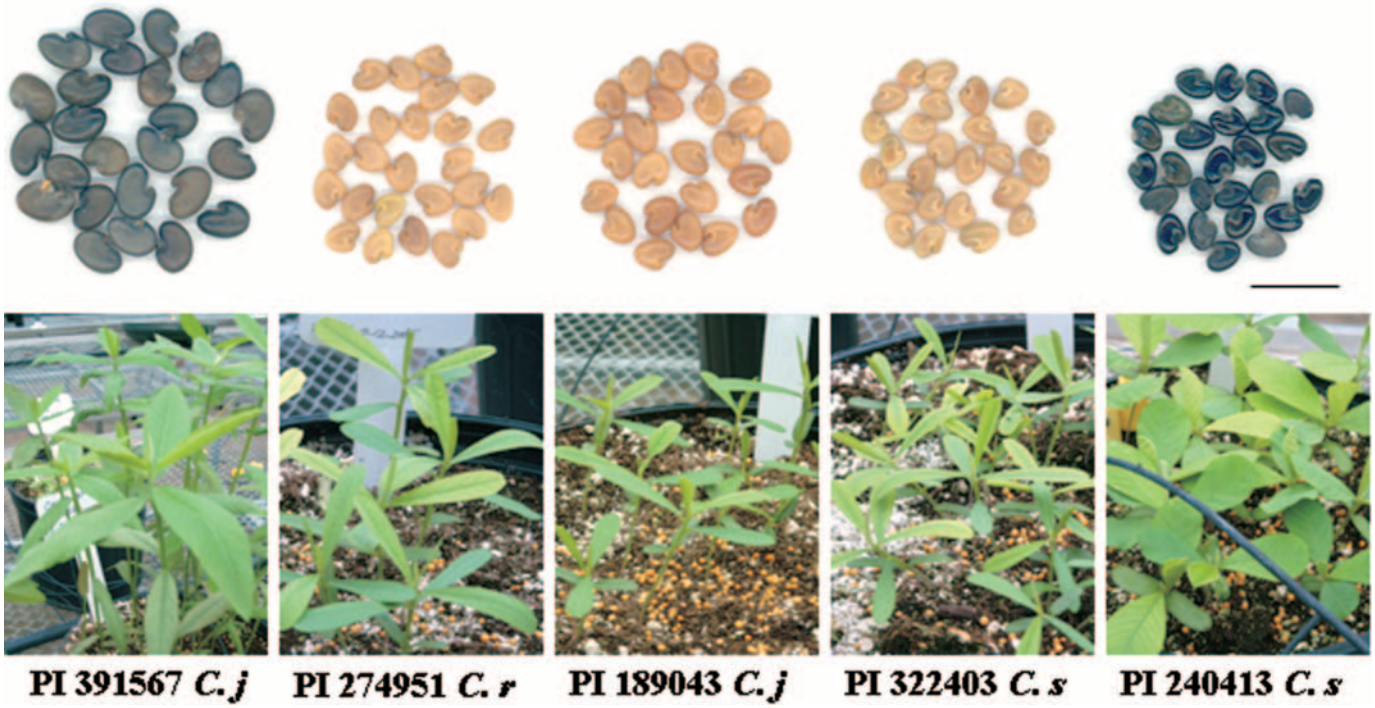


Fig. 4. Alignment of original sequence from *Medicago* with amplified sequences from *Crotalaria*. The 4 sequenced samples were *Medicago truncatula*, PI 2040413 (*Crotalaria spectabilis*), PI 468956 (*Crotalaria juncea*), and PI 207675 (*Crotalaria juncea*). The underlined sequences, bold sequences, italic sequences, and dark-shaded sequences are primers, SSRs, deleted sequences, and identical sequences, respectively.

```
Medicago   CCTCCACATAGCTGGTCGATCTCTTGATATCAAAGAGAACGAGATCCCATATAA
PI240413   CCTCCACATAGCTGGTCGATCTCTTGATATCAGGAGAGCTCGAGATCCTAGATAA
PI468956   CCTCCACATAGCTGGTCGATCTCTTGATATCAGGAGAGCTCGAGATCCTAGATAA
PI207675   CCTCCACATAGCTGGTCGATCTCTTGATATCAGGAGAGCACGAGATCCTAGATAA

Medicago   TCAAAGGATGGCCAATATAAGTCATCCGGACCAACAACTTCCCTAAGATTCTGAC
PI240413   TCAAAGGAAGGCCAATATAAGTCATTTGGGCCAACAACTTCTCTAAAGCTTTGAG
PI468956   TCAAAGGAAGGCCAATATAAGTCATTTGGGCCAACAACTTCTCTAAAGCTTTGAG
PI207675   TCAAAGGAAGGCCAATATAAGTCATTTGGGCCAACAACTTCTCTAAAGCTTTGAG

Medicago   CATTATTATTATTATTATTATTATTATCATAATTCATATTCAGAACCCTTGATGC
PI240413   C.......................ATCAGAATTG......AGAATCCGCCGTGC
PI468956   A.......................ATCGGAATTG......AGAATCCGCCGTGC
PI207675   A.......................ATCGGAATTG......AGAATCCGCCGCGT

Medicago   ATCATCATCTTTCTTTGCGTCAAACACATATAGTCCTAAACCGCACACACC
PI240413   ATCATC...TTTCACTGCATCAAATACATGTAGTCCTAAACCGCACACACC
PI468956   ATCATC...TTTCACTGCATCAAACACATGTAGTCCTAAACCGCACACACC
PI207675   ATCATC...TTTCACTGCATCAAACACATGTAGTCCTAAACCGCACACACC
```

might suggest that even though *Crotalaria* and *Medicago* diverged a long time ago, the functionality of this gene is conserved among species. Deletion and insertion around the microsatellite region has also been found in mammals. Expanded repeats in coding regions of genes are responsible for diseases such as Huntington's disease and the spinocerebellar ataxias (Zoghbi and Orr 2000). Hammock and Young (2005) found that in the regulatory region of *avpr1a*, 360 bp in and around the microsatellite region was deleted in chimpanzees, although the flanking regions were >96% conserved in human and chimpanzees. Deleting and inserting the region containing microsatellites (or SSRs) might play an important role in speciation or gene functionality diversification during the evolutionary process.

The sequence content of amplicons from *C. juncea* and *C. spectabilis* were almost identical except for a few base

differences (Fig. 4). For phylogenetic analysis, the amplicons were scored based on their sizes on agarose gels. Obviously, there was some limitation in scoring accessions with the same size of amplicons on the agarose gel due to the occurrence of point mutations. Accessions PI 468956 and PI 207675 from *C. juncea* and PI 240413 from *C. spectabilis* could not be separated from each other on an agarose gel by marker AL 365892 (Fig. 1), but they could be clearly distinguished by sequencing the amplicons derived from the same marker. In our case, more cross-species amplicons from different transferred SSR markers should be sequenced. The distinguishing power for characterization and evaluation of plant germplasm could definitely be increased by sequencing more amplicons.

Medicago and soybean are being used as research models for cool- and warm-season legumes, respectively. A vast amount of genomic information has been generated for these 2 model species. In contrast, there is not much genomic information presently available for minor legume species (such as *Crotalaria*). However, many useful traits and sources of superior alleles (genetic diversity) for agricultural production exist in the minor legume species. Genomic information (for example, DNA markers) is required to mine superior alleles or exploit the genetic diversity from germplasm. Development of DNA markers is expensive and time consuming. The transfer of EST-SSRs or genomic SSRs from model species or crops to minor species is a very efficient approach for DNA marker development. In the previous reports, transferred genomic- and EST-SSRs were used for assessment of genetic diversity of minor species in the grass family (Barkley et al. 2005; Wang et al. 2006). The present report showed the value of using genetic information from well-exploited legume models for assessing the genetic diversity of minor species of the legume family.

Acknowledgements

The authors gratefully thank Dr. Les Goertezn, Dr. Baozhu Guo, and Dr. Noelle Barkley for reviewing this manuscript, Dr. Zhenbang Chen for helping with data analysis, Ms. Lee Ann Chalkley for providing the seeds, and Ms. Meredith Reed for assisting in sample collection and DNA extraction.

References

Avise, J.C. 1994. Molecular markers, natural history and evolution. Chapman & Hall, New York, N.Y.

Barkley, N.A., Newman, M.L., Wang, M.L., Hotchkiss, M.W., and Pederson, G.A. 2005. Assessment of the genetic diversity and phylogenetic relationships of a temperate bamboo collection by using transferred EST-SSR markers. Genome, **48**: 731–737.

Chen, X., Cho, Y.G., and McCouch, S.R. 2002. Sequence divergence of rice microsatellites in *Oryza* and other plant species. Mol. Genet. Genomics, **268**: 331–343.

Dempsey, J.M. 1975. Fiber crop. The University Press of Florida, Gainesville, Fla.

Felsenstein, J. 1993. PHYLIP (phylogeny inference package). Version 3.6 [computer program]. Department of Genetics, University of Washington, Seattle, Wash.

Hammock, E.A.D., and Young, L.J. 2005. Microsatellite instability generates diversity in brain and sociobehavioral traits. Science (Washington, D.C.). **308**: 1630–1634.

Ji, X., Khan, I., Mosjidis, J.A., Wang, H., and Livant P. 2005. Variability for the presence of pyrrolizidine alkaloids in *Crotalaria juncea* L. Pharmazie, **60**: 620–622.

Miller, R.H. 1967. Crotalaria seed, morphology, anatomy, and identification. Agricultural Research Service, United States Department of Agriculture. Technical Bulletin No. 1373.

Minch, E., Ruiz-Linares, A., Goldstein, D., Feldman, M., and Cavalli-Sforza, L.L. 1996. Microsat (v. 1.5b): a computer program for calculating various statistics on microsatellite allele data. Available from http://hpgl.stanford.edu/projects/microsat/.

Morris, J.B., and Kays, S.E. 2005. Total dietary fiber variability in a cross section of *Crotalaria juncea* genetic resources. Crop Sci. **45**: 1826–1829.

Page, R.D.M. 1996. TREEVIEW. Version 1.6.6. An application to display phylogenetic trees on personal computers. Computer Appl. Biosci. **12**: 357–358.

Pandey, V.N., and Srivastava, A.K. 1990. Seed protein yield from some *Crotalaria* spp. and in vitro nutritional quality of that from *C. juncea*. Plant Foods Human Nutr. **40**: 195–200.

Purseglove, J.W. 1981. Leguminosae. *In* Tropical crops: dicotyledons. *Edited by* J.W. Purseglove. Longman Group Ltd., Essex, UK. pp. 250–254.

Ramos, M.G., Villatoro, M.A.A., Urquiaga, S., Alves, B.J.R., and Boddey, R.M. 2001. Quantification of the contribution of biological nitrogen fixation to tropical green manure crops and residual benefit to a subsequent maize crop using ^{15}N-isotope techniques. J. Biotechnol. **91**: 105–115.

Smith, J.W., and Culvenor, C.C.C. 1981. Plant sources of hepatoxic pyrrolizidinealkaloids. J. Nat. Prod. **44**: 129–152.

van Wyk, B.-E., and Schutte, A.L. 1995. Phylogenetic relationships in the tribes Podalyrieae, Liparieae and Crotalarieae. *In* Advances in legume systematics. *Edited by* M. Crisp and J.J. Doyle. Royal Botanic Gardens, Kew, UK. pp. 283–308.

Wang, M.L., Gillaspie, A.G., Newman, M.L., Dean, R.E., Pittman, R.N., Morris, J.B., and Pederson, G.A. 2004. Transfer of simple sequence repeat (SSR) markers across the legume family for germplasm characterization and evaluation. Plant Genet. Res. **2**: 107–119.

Wang, M.L., Barkley, N.A., Yu, J.K., Dean, R.E., Newman, M.L., Sorrells, M.E., and Pederson, G.A. 2005. Transfer of simple sequence repeat (SSR) markers frommajor cereal crops to minor grass species for germplasm characterization and evaluation. Plant Genet. Res. **3**: 45–57.

Wang, M.L., Chen, Z.B., Barkley, N.A., Newman, M.L., Kim, W., Raymer, P., and Pederson, G.A. 2006. Characterization of seashore paspalum (*Paspalumvaginatum* Swartz) germplasm by transferred SSRs from wheat, maize and sorghum. Genet. Res. Crop Evol. **53**: 779–791.

Williams, M.C., and Molyneux, R.J. 1987. Occurrence, concentration, and toxicity of pyrrolizidine alkaloids in *Crotalaria* seeds. Weed Sci. **36**: 476–481.

Zoghbi, H.Y., and Orr, H.T. 2000. Glutamine repeats and neurodegeneration. Annu. Rev. Neurosci. **23**: 217–247.